```
 1                          SEQUENCE LISTING
 2
 3
 4
 5   (1). GENERAL INFORMATION:
 6
 7        (i) APPLICANT(S): Kaufman, Randal J.
 8                          Wasley, Louise
 9
10       (ii) TITLE OF INVENTION:  Method of Increasing Yield of
11                                 Mature Proteins
12
13      (iii) NUMBER OF SEQUENCES:  2
14
15       (iv) CORRESPONDENCE ADDRESS:
16
17            (A) ADDRESSEE:  Genetics Institute, Inc.
18            (B) STREET:     87 CambridgePark Drive
19            (C) CITY:       Cambridge
20            (D) STATE:      Massachusetts
21            (E) COUNTRY:    United States of America
22            (F) ZIP:        02140
23
24        (v) COMPUTER READABLE FORM:
25
26            (A) MEDIUM TYPE:  Diskette, 3.50 inch, 720 Kb storage
27
28            (B) COMPUTER:  IBM PS/2
29
30            (C) OPERATING SYSTEM:  PC-DOS
31
32            (D) SOFTWARE:  WordPerfect 5.1
33
34       (vi) CURRENT APPLICATION DATA:
35
36            (A)  APPLICATION NUMBER: not yet available
37
38            (B)  FILING DATE:  26 November 1990
39
40            (C)  CLASSIFICATION:  not yet available
41
42      (vii) PRIOR APPLICATION DATA: not applicable
43
44            (A)  APPLICATION NUMBER:
45
46            (B)  FILING DATE:
47
48     (viii) ATTORNEY/AGENT INFORMATION
49
50            (A) NAME:  Ellen J. Kapinos, Esquire
51            (B) REGISTRATION NUMBER: 32,245
52            (C) REFERENCE/DOCKET NUMBER: GI 5181
53
```

54   (ix) TELECOMMUNICATION INFORMATION:
55
56   (A)   TELEPHONE: (617) 876-1170
57   (B)   TELEFAX: (617) 876-5851
58
59 (2) INFORMATION FOR SEQ ID NO:1
60
61   (i)   SEQUENCE CHARACTERISTICS:
62
63   (A)   LENGTH: 2385 base pairs
64
65   (B)   TYPE: nucleic acid
66
67   (C)   STRANDEDNESS: double
68
69   (D)   Topology: unknown
70
71   (ii) MOLECULE TYPE: partial human genomic DNA
72
73   (A) DESCRIPTION: sequence encoding furin
74
75   (iii) HYPOTHETICAL: no
76
77   (iv) ANTI-SENSE: no
78
79   (v)   PUBLICATION INFORMATION: A.M.W. van den Ouweland et al,
80   Nucl. Acids. Res., 18:664 (1990)
81
82   (vi) SEQUENCE DESCRIPTION: SEQ ID NO:1
83
84 ATG GAG CTC AGG CCC TGG TTC   21
85 Met Glu Leu Arg Pro Trp Leu
86   1   5
87
88 CTA TGG GTC GTA CCA CCA ACA GGA ACC TTG GTC CTG CTA   60
89 Leu Trp Val Val Ala Ala Thr Gly Thr Leu Val Leu Leu
90   10   15   20
91
92 GCA GCT GAT GCT CAG GGC CAG AAG GTC TTC ACC AAC ACG   99
93 Ala Ala Asp Ala Gln Gly Gln Lys Val Phe Thr Asn Thr
94   25   30
95
96 TGG GCT GTG CGC ATC CCT GGA GGC CCA GCG GTG GCC AAC   138
97 Trp Ala Val Arg Ile Pro Gly Gly Pro Ala Val Ala Asn
98   35   40   45
99
100
101 AGT GTG GCA CGG AAG CAT GGG TTC CTC AAC CTG GGC CAG   177
102 Ser Val Ala Arg Lys His Gly Phe Leu Asn Leu Gly Gln
103   50   55
104
105 ATC TTC GGG GAC TAT TAC CAC TTC TGG CAT CGA GGA GTG   216
106 Ile Phe Gly Asp Tyr Tyr His Phe Trp His Arg Gly Val

```
107   60                           65                            70
108
109
110   ACG AAG CGG TCC CTG TCG CCT CAC CGC CCG CGG CAC AGC    255
111   Thr Lys Arg Ser Leu Ser Pro His Arg Pro Arg His Ser
112           75                          80                          85
113
114
115   CGG CTG CAG AGG GAG CCT CAA GTA CAG TGG CTG GAA CAG    294
116   Arg Leu Gln Arg Glu Pro Gln Val Gln Trp Leu Glu Gln
117                       90                          95
118
119   CAG GTG GCA AAG CGA CGG ACT AAA CGG GAC GTG TAC CAG    333
120   Gln Val Ala Lys Arg Arg Thr Lys Arg Asp Val Tyr Gln
121           100                         105                         110
122
123   GAG CCC ACA GAC CCC AAG TTT CCT CAG CAG TGG TAC CTG    372
124   Glu Pro Thr Asp Pro Lys Phe Pro Gln Gln Trp Tyr Leu
125                   115                         120
126
127   TCT GGT GTC ACT CAG CGG GAC CTG AAT GTG AAG GCG GCC    411
128   Ser Gly Val Thr Gln Arg Asp Leu Asn Val Lys Ala Ala
129   125                         130                         135
130
131   TGG GCG CAG GGC TAC ACA GGG CAC GGC ATT CTG GTC TCC    450
132   Trp Ala Gln Gly Tyr Thr Gly His Gly Ile Val Val Ser
133           140                         145                         150
134
135   ATT CTG GAC GAT GGC ATC GAG AAG AAC CAC CCC GAC TTG    489
136   Ile Leu Asp Asp Gly Ile Glu Lys Asn His Pro Asp Leu
137                   155                         160
138
139   GCA GGC AAT TAT GAT CCT GGG GCC AGT TTT CAT GTC AAT    528
140   Ala Gly Asn Tyr Asp Pro Gly Ala Ser Phe Asp Val Asn
141           165                         170                         175
142
143   GAC CAG GAC CCT GAC CCC CAG CCT CGG TAC ACA CAG ATG    567
144   Asp Gln Asp Pro Asp Pro Gln Pro Arg Tyr Thr Gln Met
145                   180                         185
146
147   AAT GAC AAC AGG CAC GGC ACA CGG TGT GCG GGG GAA GTG    606
148   Asn Asp Asn Arg His Gly Thr Arg Cys Ala Gly Glu Val
149   190                         195                         200
150
151   GCT GCC GTG GCC AAC AAC CGT GTC TGT GGT GTA GGT GTG    645
152   Ala Ala Val Ala Asn Asn Gly Val Cys Gly Val Gly Val
153           205                         210                         215
154
155   GCC TAC AAC GCC CGC ATT GGA GGG GTC CGC ATG CTG GAT    684
156   Ala Tyr Asn Ala Arg Ile Gly Gly Val Arg Met Leu Asp
157                   220                         225
158
159   GGC GAG GTG ACA GAT GCA GTG GAG GCA CGC TCG CTG GGC    723
```

```
160  Gly Glu Val Thr Asp Ala Val Glu Ala Arg Ser Leu Gly
161      230             235             240
162
163  CTG AAC CCC AAC CAC ATC CAC ATC TAC AGT GCC AGC TGG     762
164  Leu Asn Pro Asn His Ile His Ile Tyr Ser Ala Ser Trp
165          245             250
166
167  GGC CCC GAG GAT GAC GGC AAG ACA GTG GAT GGG CCA GCC     801
168  Gly Pro Glu Asp Asp Gly Lys Thr Val Asp Gly Pro Ala
169  255             260             265
170
171  CGG CTC GCC GAG GAG GCC TTC TTC CGT GGG CTT AGC CAG     840
172  Arg Leu Ala Glu Glu Ala Phe Phe Arg Gly Val Ser Gln
173          270             275             280
174
175  GGC CGA GGG GGG CTG GGC TCC ATC TTT GTC TGG GCC TCG     879
176  Gly Arg Gly Gly Leu Gly Ser Ile Phe Val Trp Ala Ser
177              285             290
178
179  GGG AAC GGG GGG CGG GAA CAT GAC AGC TGC AAC TGC GAC     918
180  Gly Asn Gly Gly Arg Glu His Asp Ser Cys Asn Cys Asp
181      295             300             305
182
183  GGC TAC ACC AAC AGT ATC TAC ACG CTG TCC ATC AGC AGC     957
184  Gly Tyr Thr Asn Ser Ile Tyr Thr Leu Ser Ile Ser Ser
185              310             315
186
187  GCC ACG CAG TTT GGC AAC GTG CCG TGG TAC AGC GAG GCC     996
188  Ala Thr Gln Phe Gly Asn Val Pro Trp Tyr Ser Glu Ala
189  320             325             330
190
191  TGC TCG TCC ACA CTG GCC ACG ACC TAC AGC AGT GGC AAC    1035
192  Cys Ser Ser Thr Leu Ala Thr Thr Tyr Ser Ser Gly Asn
193          335             340             345
194
195  CAG AAT GAG AAG CAG ATC GTG ACG ACT GAC TTG CGG CAG    1074
196  Gln Asn Glu Lys Gln Ile Val Thr Thr Asp Leu Arg Gln
197              350             355
198
199  AAG TGC ACG GAG TCT CAC ACG GGC ACC TCA GCC TCT GCC    1113
200  Lys Cys Thr Glu Ser His Thr Gly Thr Ser Ala Ser Ala
201      360             365             370
202
203  CCC TTA GCA GCC GGC ATC ATT GCT CTC ACC CTG GAG GCC    1152
204  Pro Leu Ala Ala Gly Ile Ile Ala Leu Thr Leu Glu Ala
205              375             380
206
207  AAT AAG AAC CTC ACA TGG CGG GAC ATG CAA CAC CTG GTG    1191
208  Asn Lys Asn Leu Thr Trp Arg Asp Met Gln His Leu Val
209  385             390             395
210
211  GTA CAG ACC TCG AAG CCA GCC CAC CTC AAT GCC AAC GAC    1230
212  Val Gln Thr Ser Lys Pro Ala His Leu Asn Ala Asn Asp
```

```
213           400                     405                     410
214
215   TGG GCC ACC AAT GGT GTG GGG CGG AAA GTG AGC CAC TCA   1269
216   Trp Ala Thr Asn Gly Val Gly Arg Lys Val Ser His Ser
217                   415                     420
218
219   TAT GGC TAC GGG CTT TTC GAC GCA GGC GCC ATG GTG GCC   1308
220   Tyr Gly Tyr Gly Leu Leu Asp Ala Gly Ala Met Val Ala
221       425                     430                     435
222
223   CTG GCC CAG AAT TGG ACC ACA GTC GCC CCC CAG CGG AAG   1347
224   Leu Ala Gln Asn Trp Thr Thr Val Ala Pro Gln Arg Lys
225                   440                     445
226
227   TGC ATC ATC GAC ATC CTC ACC GAG CCC AAA GAC ATC GGG   1386
228   Cys Ile Ile Asp Ile Leu Thr Glu Pro Lys Asp Ile Gly
229   450                     455                     460
230
231   AAA CGG CTC GAC CTC CGG AAC ACC GTG ACC GCG TCC CTG   1425
232   Lys Arg Leu Glu Val Arg Lys Thr Val Thr Ala Cys Leu
233           465                     470                     475
234
235   GGC GAG CCC AAC CAC ATC ACT CGG CTG GAG CAC GCT CAG   1464
236   Gly Glu Pro Asn His Ile Thr Arg Leu Glu His Ala Gln
237                   480                     485
238
239   GCG CGG CTC ACC CTG TCC TAT AAT CGC CGT GGC GAC CTG   1503
240   Ala Arg Leu Thr Leu Ser Tyr Asn Arg Arg Gly Asp Leu
241           490                     495                     500
242
243   GCC ATC CAC CTG GTC AGC CCC ATG GGC ACC CGC TCC ACC   1542
244   Ala Ile His Leu Val Ser Pro Met Gly Thr Arg Ser Thr
245                   505                     510
246
247   CTG CTG GCA GCC AGG CCA CAT GAC TAC TCC GCA GAT GGG   1581
248   Leu Leu Ala Ala Arg Pro His Asp Tyr Ser Ala Asp Gly
249   515                     520                     525
250
251   TTT AAT GAC TGG GCC TTC ATG ACA ACT CAT TCC TGG GAT   1620
252   Phe Asn Asp Trp Ala Phe Met Thr Thr His Ser Trp Asp
253                   530                     535                     540
254
255   GAC GAT CCC TCT GGG GAG TGG GTC CTA GAG ATT GAA AAC   1659
256   Glu Asp Pro Ser Gly Glu Trp Val Leu Glu Ile Glu Asn
257                   545                     550
258
259   ACC AGC GAA GCC AAC AAC TAT GGG ACG CTG ACC AAC TCC   1698
260   Thr Ser Glu Ala Asn Asn Tyr Gly Thr Leu Thr Lys Phe
261           555                     560                     565
262
263   ACC CTC GTA CTC TAT GGC ACC GCC CCT GAC GGG CTC CCC   1737
264   Thr Leu Val Leu Tyr Gly Thr Ala Pro Glu Gly Leu Pro
265           570                     575
```

```
266
267 GTA CCT CCA GAA AGC AGT GGC TGC AAG ACC CTC ACG TCC  1776
268 Val Pro Pro Glu Ser Ser Gly Cys Lys Thr Leu Thr Ser
269 580                 585                 590
270
271 AGT CAG GCC TGT GTG GTG TGC GAG GAA GGC TTC TCC CTC  1815
272 Ser Gln Ala Cys Val Val Cys Glu Glu Gly Phe Ser Leu
273         595                 600                 605
274
275 CAC CAG AAG AGC TGT GTC CAG CAC TGC CCT CCA GGC TTC  1854
276 His Gln Lys Ser Cys Val Gln Asn Cys Pro Pro Gly Phe
277                 610                 615
278
279 GCC CCC CAA GTC CTC GAT ACG CAC TAT AGC ACC GAG AAT  1893
280 Ala Pro Gln Val Leu Asp Thr Asn Tyr Ser Thr Glu Asn
281     620                 625                 630
282
283 GAC GTG GAG ACC ATC CGG GCC AGC GTC TGC GCC CCC TGC  1932
284 Asp Val Glu Thr Ile Arg Ala Ser Val Cys Ala Pro Cys
285             635                 640
286
287 CAC GCC TCA TGT GCC ACA TGC CAG GGG CCG GCC CTG ACA  1971
288 His Ala Ser Cys Ala Thr Cys Gln Gly Pro Ala Leu Thr
289 645                 650                 655
290
291 GAC TGC CTC AGC TGC CCC AGC CAC GCC TCC TTG GAC CCT  2010
292 Asp Cys Leu Ser Cys Pro Ser His Ala Ser Leu Asp Pro
293             660                 665                 670
294
295 GTG GAG CAG ACT TGC TCC CGG CAA AGC CAG AGC AGC CGA  2049
296 Val Glu Gln Thr Cys Ser Arg Gln Ser Gln Ser Ser Arg
297             675                 680
298
299 GAG TCC CCG CCA CAG CAG CAG CCA CCT CGG CTG CCC CCG  2088
300 Glu Ser Pro Pro Gln Gln Gln Pro Pro Arg Leu Pro Pro
301     685                 690                 695
302
303 GAG GTG GAG GCG GGG CAA CGG CTG CGG GCA GGG CTG CTG  2127
304 Glu Val Glu Ala Gly Gln Arg Leu Arg Ala Gly Leu Leu
305             700                 705
306
307 CCC TCA CAC CTG CCT GAG GTG GTG GCC GGC CTC AGC TGC  2166
308 Pro Ser His Leu Pro Glu Val Val Ala Gly Leu Ser Cys
309 710                 715                 720
310
311 GCC TTC ATC GTG CTG GTC TTC GTC ACT GTC TTC CTG CTC  2205
312 Ala Phe Ile Val Leu Val Phe Val Thr Val Phe Leu Val
313             725                 730                 735
314
315 CTG CAG CTG CGC TCT GGC TTT AGT TTT CGG GGG GTG AAG  2244
316 Leu Gln Leu Arg Ser Gly Phe Ser Phe Arg Gly Val Lys
317                 740                 745
318
```

```
319  GTG TAC ACC ATG GAC CGT GGC CTC ATC TCC TAC AAG GGG   2283
320  Val Tyr Thr Met Asp Arg Gly Leu Ile Ser Tyr Lys Gly
321      750                 755                 760
322
323  CTG CCC CCT GAA GCC TGG CAG GAG GAG TGC CCG TCT GAC   2322
324  Leu Pro Pro Glu Ala Trp Gln Glu Glu Cys Pro Ser Asp
325              765                 770
326
327  TCA GAA GAG GAC GAG GGC CGG GGC GAG AGG ACC GCC TTT   2361
328  Ser Glu Glu Asp Glu Gly Arg Gly Glu Arg Thr Ala Phe
329  775                 780                 785
330
331  ATC AAA GAC CAG AGC GCC CTC TGA                       2385
332  Ile Lys Asp Gln Ser Ala Leu End
333          790
334
335
336  (3)   INFORMATION FOR SEQ ID NO:2
337
338        (i)   SEQUENCE CHARACTERISTICS:
339
340              (A)   LENGTH: 794 amino acids
341
342              (B)   TYPE:   amino acids
343
344              (C)   STRANDEDNESS:   single
345
346              (D)   Topology:   unknown
347
348        (ii)  MOLECULE TYPE: furin
349
350       (iii)  HYPOTHETICAL:   no
351
352        (iv)  ANTI-SENSE:   no
353
354        (v)   PUBLICATION INFORMATION: A.M.W. van den Ouweland et al,
355                                   Nucl. Acids. Res., 18:664 (1990)
356
357        (vi)  SEQUENCE DESCRIPTION:   SEQ ID NO:2
358
359
360  Met Glu Leu Arg Pro Trp Leu
361    1               5
362
363  Leu Trp Val Val Ala Ala Thr Gly Thr Leu Val Leu Leu
364          10                  15                  20
365
366  Ala Ala Asp Ala Gln Gly Gln Lys Val Phe Thr Asn Thr
367                  25                  30
368
369  Trp Ala Val Arg Ile Pro Gly Gly Pro Ala Val Ala Asn
370          35                  40                  45
371
```

```
372   Ser Val Ala Arg Lys His Gly Phe Leu Asn Leu Gly Gln
373               50              55
374
375   Ile Phe Gly Asp Tyr Tyr His Phe Trp His Arg Gly Val
376    60              65              70
377
378   Thr Lys Arg Ser Leu Ser Pro His Arg Pro Arg His Ser
379           75              80                      85
380
381
382   Arg Leu Gln Arg Glu Pro Gln Val Gln Trp Leu Glu Gln
383               90              95
384
385   Gln Val Ala Lys Arg Arg Thr Lys Arg Asp Val Tyr Gln
386       100             105             110
387
388   Glu Pro Thr Asp Pro Lys Phe Pro Gln Gln Trp Tyr Leu
389           115             120
390
391   Ser Gly Val Thr Gln Arg Asp Leu Asn Val Lys Ala Ala
392   125             130             135
393
394   Trp Ala Gln Gly Tyr Thr Gly His Gly Ile Val Val Ser
395       140             145                     150
396
397   Ile Leu Asp Asp Gly Ile Glu Lys Asn His Pro Asp Leu
398           155             160
399
400   Ala Gly Asn Tyr Asp Pro Gly Ala Ser Phe Asp Val Asn
401       165             170             175
402
403   Asp Gln Asp Pro Asp Pro Gln Pro Arg Tyr Thr Gln Met
404           180             185
405
406   Asn Asp Asn Arg His Gly Thr Arg Cys Ala Gly Glu Val
407   190             195             200
408
409   Ala Ala Val Ala Asn Asn Gly Val Cys Gly Val Gly Val
410       205             210             215
411
412   Ala Tyr Asn Ala Arg Ile Gly Gly Val Arg Met Leu Asp
413               220             225
414
415   Gly Glu Val Thr Asp Ala Val Glu Ala Arg Ser Leu Gly
416       230             235             240
417
418   Leu Asn Pro Asn His Ile His Ile Tyr Ser Ala Ser Trp
419           245             250
420
421   Gly Pro Glu Asp Asp Gly Lys Thr Val Asp Gly Pro Ala
422   255             260             265
423
424   Arg Leu Ala Glu Glu Ala Phe Phe Arg Gly Val Ser Gln
```

```
425              270                    275                    280
426
427   Gly Arg Gly Gly Leu Gly Ser Ile Phe Val Trp Ala Ser
428                   285                    290
429
430   Gly Asn Gly Gly Arg Glu His Asp Ser Cys Asn Cys Asp
431       295                300                305
432
433   Gly Tyr Thr Asn Ser Ile Tyr Thr Leu Ser Ile Ser Ser
434                   310                315
435
436
437   Ala Thr Gln Phe Gly Asn Val Pro Trp Tyr Ser Glu Ala
438   320                325                330
439
440   Cys Ser Ser Thr Leu Ala Thr Thr Tyr Ser Ser Gly Asn
441           335                340                345
442
443   Gln Asn Glu Lys Gln Ile Val Thr Thr Asp Leu Arg Gln
444                   350                355
445
446   Lys Cys Thr Glu Ser His Thr Gly Thr Ser Ala Ser Ala
447           360                365                370
448
449   Pro Leu Ala Ala Gly Ile Ile Ala Leu Thr Leu Glu Ala
450               375                380
451
452   Asn Lys Asn Leu Thr Trp Arg Asp Met Gln His Leu Val
453   385                390                395
454
455   Val Gln Thr Ser Lys Pro Ala His Leu Asn Ala Asn Asp
456           400                405                410
457
458   Trp Ala Thr Asn Gly Val Gly Arg Lys Val Ser His Ser
459                   415                420
460
461   Tyr Gly Tyr Gly Leu Leu Asp Ala Gly Ala Met Val Ala
462           425                430                435
463
464   Leu Ala Gln Asn Trp Thr Thr Val Ala Pro Gln Arg Lys
465               440                445
466
467   Cys Ile Ile Asp Ile Leu Thr Glu Pro Lys Asp Ile Gly
468   450                455                460
469
470   Lys Arg Leu Glu Val Arg Lys Thr Val Thr Ala Cys Leu
471           465                470                475
472
473   Gly Glu Pro Asn His Ile Thr Arg Leu Glu His Ala Gln
474                   480                485
475
476   Ala Arg Leu Thr Leu Ser Tyr Asn Arg Arg Gly Asp Leu
477           490                495                500
```

```
478
479   Ala Ile His Leu Val Ser Pro Met Gly Thr Arg Ser Thr
480           505             510
481
482   Leu Leu Ala Ala Arg Pro His Asp Tyr Ser Ala Asp Gly
483   515             520             525
484
485   Phe Asn Asp Trp Ala Phe Met Thr Thr His Ser Trp Asp
486       530             535             540
487
488   Glu Asp Pro Ser Gly Glu Trp Val Leu Glu Ile Glu Asn
489               545             550
490
491   Thr Ser Glu Ala Asn Asn Tyr Gly Thr Leu Thr Lys Phe
492       555             560             565
493
494   Thr Leu Val Leu Tyr Gly Thr Ala Pro Glu Gly Leu Pro
495           570             575
496
497   Val Pro Pro Glu Ser Ser Gly Cys Lys Thr Leu Thr Ser
498   580             585             590
499
500   Ser Gln Ala Cys Val Val Cys Glu Glu Gly Phe Ser Leu
501       595             600             605
502
503   His Gln Lys Ser Cys Val Gln Asn Cys Pro Pro Gly Phe
504           610             615
505
506   Ala Pro Gln Val Leu Asp Thr Asn Tyr Ser Thr Glu Asn
507       620             625             630
508
509   Asp Val Glu Thr Ile Arg Ala Ser Val Cys Ala Pro Cys
510           635             640
511
512   His Ala Ser Cys Ala Thr Cys Gln Gly Pro Ala Leu Thr
513   645             650             655
514
515   Asp Cys Leu Ser Cys Pro Ser His Ala Ser Leu Asp Pro
516       660             665             670
517
518   Val Glu Gln Thr Cys Ser Arg Gln Ser Gln Ser Ser Arg
519           675             680
520
521   Glu Ser Pro Pro Gln Gln Gln Pro Pro Arg Leu Pro Pro
522       685             690             695
523
524   Glu Val Glu Ala Gly Gln Arg Leu Arg Ala Gly Leu Leu
525           700             705
526
527   Pro Ser His Leu Pro Glu Val Val Ala Gly Leu Ser Cys
528   710             715             720
529
530   Ala Phe Ile Val Leu Val Phe Val Thr Val Phe Leu Val
```

```
531              725                    730                    735
532
533  Leu Gln Leu Arg Ser Gly Phe Ser Phe Arg Gly Val Lys
534              740                    745
535
536  Val Tyr Thr Met Asp Arg Gly Leu Ile Ser Tyr Lys Gly
537      750                  755                760
538
539  Leu Pro Pro Glu Ala Trp Gln Glu Glu Cys Pro Ser Asp
540              765                    770
541
542  Ser Glu Glu Asp Glu Gly Arg Gly Glu Arg Thr Ala Phe
543  775                  780                785
544
545
546  Ile Lys Asp Gln Ser Ala Leu End
547              790
```

remove (handwritten annotation next to circled "End")

Line	Error	Original text
5	Unknown or Misplaced Identifier	(1). GENERAL INFORMATION:
7	Unknown or Misplaced Identifier	(i) APPLICANT(S): Kaufman, Randal J.
8	Unknown or Misplaced Identifier	Wasley, Louise
10	Unknown or Misplaced Identifier	(ii) TITLE OF INVENTION: Method of Increasing
11	Unknown or Misplaced Identifier	Mature Proteins
13	Unknown or Misplaced Identifier	(iii) NUMBER OF SEQUENCES: 2
15	Unknown or Misplaced Identifier	(iv) CORRESPONDENCE ADDRESS:
17	Unknown or Misplaced Identifier	(A) ADDRESSEE: Genetics Institute, Inc.
18	Unknown or Misplaced Identifier	(B) STREET: 87 CambridgePark Drive
19	Unknown or Misplaced Identifier	(C) CITY: Cambridge
20	Unknown or Misplaced Identifier	(D) STATE: Massachusetts
21	Unknown or Misplaced Identifier	(E) COUNTRY: United States of America
22	Unknown or Misplaced Identifier	(F) ZIP: 02140
24	Unknown or Misplaced Identifier	(v) COMPUTER READABLE FORM:
26	Unknown or Misplaced Identifier	(A) MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb
28	Unknown or Misplaced Identifier	(B) COMPUTER: IBM PS/2
30	Unknown or Misplaced Identifier	(C) OPERATING SYSTEM: PC-DOS
32	Unknown or Misplaced Identifier	(D) SOFTWARE: WordPerfect 5.1
34	Unknown or Misplaced Identifier	(vi) CURRENT APPLICATION DATA:
36	Unknown or Misplaced Identifier	(A) APPLICATION NUMBER: not yet available
38	Unknown or Misplaced Identifier	(B) FILING DATE: 26 November 1990
40	Unknown or Misplaced Identifier	(C) CLASSIFICATION: not yet available
42	Unknown or Misplaced Identifier	(vii) PRIOR APPLICATION DATA: not applicable
44	Unknown or Misplaced Identifier	(A) APPLICATION NUMBER:
46	Unknown or Misplaced Identifier	(B) FILING DATE:
48	Unknown or Misplaced Identifier	(viii) ATTORNEY/AGENT INFORMATION
50	Unknown or Misplaced Identifier	(A) NAME: Ellen J. Kapinos, Esquire
51	Unknown or Misplaced Identifier	(B) REGISTRATION NUMBER: 32,245
52	Unknown or Misplaced Identifier	(C) REFERENCE/DOCKET NUMBER: GI 5181
54	Unknown or Misplaced Identifier	(ix) TELECOMMUNICATION INFORMATION:
56	Unknown or Misplaced Identifier	(A) TELEPHONE: (617) 876-1170
57	Unknown or Misplaced Identifier	(B) TELEFAX: (617) 876-5851
80	Extra Level-0 Records	Nucl. Acids. Res., 18:664 (1990)
355	Extra Level-0 Records	Nucl. Acids. Res., 18:664 (1990)
546	Wrong Amino Acid Designator	Ile Lys Asp Gln Ser Ala Leu End

Patent Application US 07/621,092

Manditory Identifier that was not found
GENERAL INFORMATION
APPLICANT
TITLE OF INVENTION
NUMBER OF SEQUENCES
CORRESPONDENCE ADDRESS
ADDRESSEE
STREET
CITY
STATE
COUNTRY
ZIP
COMPUTER READABLE FORM
MEDIUM TYPE
COMPUTER
OPERATING SYSTEM
SOFTWARE
CURRENT APPLICATION DATA
APPLICATION NUMBER
FILING DATE
CLASSIFICATION
PRIOR APPLICATION DATA
APPLICATION NUMBER
FILING DATE